


Do Men and Women Exhibit Different Preferences for Mates? A Replication of Eastwick and Finkel (2008)

SAGE Open
July-September 2015: 1–14
© The Author(s) 2015
DOI: 10.1177/2158244015605160
sgo.sagepub.com


Dylan F. Selterman¹, Elizabeth Chagnon¹,
and Sean P. Mackinnon²

Abstract

Evolutionary theory predicts that men will prefer physically attractive romantic partners, and women will prefer wealthy, high-status partners. This theory is well-supported when examining ideal hypothetical partner preferences, but less support has been found when people interact face-to-face. The present study served as a direct replication of results reported in Eastwick and Finkel (2008). We recruited 307 participants and utilized a speed-dating methodology to allow in-person interactions, then administered follow-up surveys to measure romantic interest over 30 days. Data were analyzed using multilevel modeling and were aggregated using meta-analysis. Consistent with previous findings, our results showed that participants were more romantically interested in potential partners if they were viewed as attractive and good potential earners, and these associations were *not* moderated by gender. Results suggest that gender differences predicted by evolutionary theory may not hold when people interact with potential romantic partners face-to-face. However, we discuss these results in light of some general methodological limitations and evidence from other lines of research.

Keywords

speed dating, sex differences, attraction, evolutionary theory, replication

Traditional evolutionary psychological theory, and more specifically, Sexual Strategies Theory (SST; Buss & Schmitt, 1993), stipulates that men and women face different challenges to reproductive success (e.g., paternity uncertainty) and therefore would prioritize qualities differently in romantic mates. In a heterosexual mating context, men (more so than women) are theorized to favor young partners with higher levels of physical attractiveness, because youth and attractiveness convey information about her reproductive fitness (increased likelihood of bearing healthy children for the male). In contrast, women (more so than men) are thought to favor partners with higher levels of social status and economic/material resources, because a stronger commitment of resources from a male partner will increase the odds that her children will survive to adulthood. These hypotheses rely greatly on Trivers's (1972) parental investment theory, which proposes that women are more selective in their mates because the amount of investment they must put into caring for offspring has been much greater throughout our evolutionary history. Fletcher, Simpson, Thomas, and Giles (1999) found that a three-dimensional model of hypothetical romantic partner ideals captured the most salient features of ideal romantic partners: warmth-trustworthiness, vitality-attractiveness, and status-resources.

Research using hypothetical mate preference surveys has tended to support evolutionary theory, finding that (a) men place higher value on vitality-attractiveness, (b) women place higher value on status-resources, and (c) men and women value warmth-trustworthiness about equally (Buss, Shackelford, Kirkpatrick, & Larsen, 2001; Feingold, 1992).

However, in a face-to-face setting, it is not clear that men and women would display such differences in the live experience of romantic attraction. Indeed, some studies have suggested that the gendered mate preferences predicted by evolutionary theory might not hold when potential mates actually interact with each other (Eastwick & Finkel, 2008; Kurzban & Weeden, 2005; Luo & Zhang, 2009). However, some other recent studies utilizing live-attraction methodology have found some significant gender interactions in the direction predicted by related evolutionary theories. Among other findings, those studies showed women being more

¹University of Maryland, College Park, USA

²Dalhousie University, Halifax, Nova Scotia, Canada

Corresponding Author:

Dylan F. Selterman, Department of Psychology, University of Maryland, College Park, MD 20742, USA.

Email: dselter@umd.edu



choosy/selective than men (Fletcher, Kerr, Li, & Valentine, 2014), women valuing earning prospects and men valuing attractiveness more than their counterparts (Li et al., 2013), women valuing intelligence more than men (Fisman, Iyengar, Kamenica, & Simonson, 2006), and men valuing attractiveness more than women (Fisman et al., 2006). Notably, a recent meta-analysis compiling live-attraction data from 97 studies (Eastwick, Luchies, Finkel, & Hunt, 2014) also suggested these gendered mate preferences do not hold when analyzed in aggregate. Based on the studies to date utilizing live-attraction methods, the findings are mixed with meta-analytic evidence leaning toward no gender differences in short-term, face-to-face dating contexts.

The current study was designed to assess potential similarities and differences between men and women in terms of mate preferences (what features they would seek in an ideal romantic partner), in a live interaction context, rather than hypothetical preference surveys (as has been typically done in the bulk of previous research). Our main goal was to replicate the key results reported in Eastwick and Finkel's (2008) work. The present study represents a small portion of the Reproducibility Project, a collaborative, large-scale project that is attempting to determine the reproducibility of psychological research (Open Science Collaboration, 2014). In this project, teams of researchers attempt to independently replicate key findings from psychological studies selected from three high-profile psychology journals published in 2008 (*Journal of Personality and Social Psychology*, *Psychological Science*, *Journal of Experimental Psychology: Learning, Memory, and Cognition*). The present results represent a replication attempt of Eastwick and Finkel's (2008) article on gender differences in mate preferences. Although Eastwick and Finkel (2008) tested numerous hypotheses, the present study attempted to replicate core findings regarding how gender did not moderate associations between potential mate characteristics (i.e., physical attractiveness, earning prospects, and personal warmth) and romantic interest. Consistent with Eastwick and Finkel's (2008) findings, we hypothesized the following:

Hypothesis 1: Perceived earning prospects (X1), physical attractiveness (X2), and warmth (X3) in potential partners will all positively predict romantic interest (Y) in a live-attraction setting, for both men and women.

Hypothesis 2: Contrary to SST, the degree to which perceived attractiveness, earning prospects, and warmth all predict romantic interest will *not* be moderated by gender. Put another way, on average, everyone (regardless of gender) will tend to prefer partners who are perceived to have these qualities to the same extent.

For clarity, a significant gender *difference* in our study (meaning, for example, that for women, a potential partner's earning prospects was a stronger predictor of feelings of attraction than it was for men) would indicate a failed

replication. Although the original study contains many effects, the "key result" we attempted to replicate for the Reproducibility Project was the non-significant, two-way interaction between gender (moderator) and earning prospects (X2) in predicting romantic interest (Y). However, in the present report, we also test similar hypotheses for physical attractiveness and how personable a person is (i.e., warmth) in the interest of providing a more comprehensive replication of Eastwick and Finkel (2008).

Method

Open Materials and Data

All the materials (i.e., questionnaires), research protocols, raw data files, and statistical syntax files are archived online using the Open Science Framework website. In addition, this website records the pre-registered nature of this replication as well as draft copies of this manuscript. All these materials can be accessed at <https://osf.io/ng6cc>

Power Analysis

Because the core finding of the original Eastwick and Finkel (2008) study was a null effect (i.e., a non-significant interaction effect), we needed to arbitrarily choose an effect size that we deemed practically significant in this context. For this study, we settled on a medium-small effect size ($r = .20$), based on Cohen's (1988) guidelines. Moreover, in a multi-level model, power depends both on the number of participants (Level 2) and the number of matches per participant (Level 1).¹ Eastwick and Finkel (2008) reported 2.5 matches per participant, on average, so we used this value. Power analysis for this multilevel model was conducted using PINT software (Bosker, Snijders, & Guldmond, 2003). Assuming an alpha of .05, an effect size of $r = .20$, 2.5 participants in each Level 2 group, no correlation between the predictor and moderator, Level 1 residual variance of .91, and intercept variance of .06, this power analysis suggested that we will require 275 participants to have 80% power, 380 participants to have 90% power, and 480 participants to have 95% power. Given the difficulty of recruiting participants from a semi-specialized sample using complex, time-consuming methodology, we decided on a planned sample size of 275 (80% power).

Participants

The sample consisted of undergraduate students at the University of Maryland. The only pre-selection rule was that participants be single (i.e., not in a romantic relationship) and at least 18 years of age. Participants were recruited via the psychology department subject pool and word of mouth. Overall, 307 people (49.8% female) participated in this study, and each person had average of 9.28 ($SD = 2.14$) speed

dates. This is somewhat larger than our planned sample size of 275; this is because we left sign ups open for all interested students to participate in a speed-dating event during the Fall 2012 and Spring 2013 semesters. Only 141 (45.9%) participants completed the age and ethnicity demographic variables; the demographic survey was given at the completion of the study, and there was high attrition rate following the follow-up surveys (we did not utilize a pretesting questionnaire). These participants were an average of 18.99 years old ($SD = 2.86$). The sample primarily identified as White (62.4%), Hispanic (9.9%), African American (10.6%), Asian/Pacific Islander (11.3%), or “Other” (5.7%).

Materials

Materials used in the present study were identical to those used in Eastwick and Finkel (2008). All measures included in the present report (except for “yessing”) were assessed on four different questionnaires: pre-event, interaction record, post-match, and follow-up. The present report focuses on the following three traits that might describe a romantic partner: *physically attractive* (assessed by the items “physically attractive” and “sexy/hot”), *earning prospects* (“good career prospects,” “ambitious/driven”), and *personable* (“fun/exciting,” “responsive,” “dependable/trustworthy,” “friendly/nice”).

As part of the interaction record, participants rated on a scale from 1 (*not at all*) to 9 (*extremely*) the extent to which they thought each speed-dating partner was characterized by the items reported above that assessed physically attractive, earning prospects, and personable characteristics.

On a scale from 1 (*strongly disagree*) to 9 (*strongly agree*), the interaction record assessed participants’ reports of *romantic desire* for the speed-dating partner (“I really liked my interaction partner,” “I was sexually attracted to my interaction partner,” and “I am likely to say ‘yes’ to my interaction partner”) and *chemistry* with him or her (“My interaction partner and I seemed to have a lot in common,” “. . . seemed to have similar personalities,” and “. . . had a real connection”).

Procedure

A “speed-dating” paradigm was used (see Finkel, Eastwick, & Matthews, 2007), in which participants went on brief, 4-min “dates” with potential romantic partners. Prior to the event, participants completed a pre-event questionnaire, which assessed their ideal romantic partner traits and how much they desired a “serious” relationship. Eighteen speed-dating events were held, with 14 to 30 total participants (7–15 men and women) per event. After the speed-dating event, participants were given the opportunity to select people they were romantically interested in. If two participants mutually “matched” with each other, the research team gave both participants each other’s contact information (email addresses),

to contact each other as they saw fit. A post-match survey measured the degree to which they liked their matches. We did not monitor the communication between participants, but instead relied on participants to self-report the status of their interactions during each follow-up survey. We assessed participants’ degree of attraction (see “Materials” section) and relationship initiation with follow-up questions that were dependent on whether or not participants had subsequent interactions with mutual matches (e.g., “What is the current status of your relationship with [name]?” [a] dating seriously, [b] dating casually, etc.). These “pivot” questions were administered only if the participant and match “hung out” again, engaged in romantic behavior including dates and/or labeled the match *a romantic partner* (or having romantic potential), engaged in sexual activity, and so on. We administered follow-up surveys to all participants with at least one mutual match; we emailed surveys about their match(es) to participants once every 3 days for 30 days in total (10 follow-up surveys per participant). If a participant had indicated no romantic relationship with a match for 3 or more follow-up surveys, we removed that match from their survey to reduce survey fatigue. Participants were also given the opportunity to “write-in” anyone else they were dating during this period if they met elsewhere beyond the initial speed-dating event.² All participants who did not have a match from the speed-dating event were sent a generic follow-up survey that only asked whether they had a person to “write in” to the study. During the 10th and final follow-up survey, we also assessed demographic information as well as revisited the pre-event questionnaires items on ideal partner preferences. To maintain the integrity of the original study being replicated, any additional demographics items were administered at the very end of the study. Low response rates to the demographic items in the final questionnaire may be due to participant survey fatigue.

The full survey materials are provided separately (see <https://osf.io/ng6cc/>).

Data Analytic Plan

Eastwick and Finkel (2008) analyzed data using SAS® statistical software. After examining syntax provided by the original authors, it was apparent that data could be analyzed in a comparable way using different statistical software. Because the research team had more experience using Mplus software (Muthén & Muthén, 2012), hypothesis testing was conducted using Mplus 7.0 software. Missing data were handled using a full information maximum likelihood method. Gender was contrast coded -0.5 (male) and $+0.5$ (female). All other variables were left unstandardized as summed scale totals. All items in multi-item scales were retained in the original format, even if reliability assessed with Cronbach’s alpha was low. In multilevel analyses, MLR (Maximum Likelihood Robust) estimation was used (i.e., an alternative estimate of standard errors, which is more robust to

violations of the normality assumption) instead of ML (Maximum Likelihood) estimation, which assumes multivariate normality.

The first step in analyses was to conduct all the two-level models (matches nested within participants). The predictor variables were gender (Level 2), perceived earning prospects in dates/matches (Level 1), and the interaction effect (i.e., the multiplicative product of gender and earning prospects). There were six dependent variables measured at Level 1 being considered in the two-level models: romantic desire, chemistry, “yessing,”³ excitement, initiation plans, and initiation hopes. Thus, we conducted six separate analyses, one for each dependent variable. We specified random intercepts, but fixed slopes (i.e., the intercepts are allowed to freely vary across participants, but the slopes are constrained to equality across participants).

The second step in the analysis was to conduct all the three-level models (Level 1 = repeated measures of romantic interest; Level 2 = matches; Level 3 = participants). Again, the predictor variables were gender (Level 3), perceived earning prospects (Level 2), and the interaction effect. There were eight dependent variables measured at Level 1 being considered in these three-level models: desire to get to know the other person better, date initiation, date enjoyment, passion, desire for a one-night stand, desire for a casual relationship, desire for a serious relationship, and commitment. Thus, we conducted eight separate analyses, one for each dependent variable. Again, we specified intercepts as random, but all slopes as fixed, consistent with Eastwick and Finkel (2008).

In the final step, all 14 analyses were combined using meta-analysis to examine whether there was an overall gender difference across all the analyses. First, we calculated an unstandardized regression coefficient for the relationship between earning prospects and outcomes for men and women separately (based on the 14 analyses above). This resulted in 28 unstandardized coefficients in total. Following this step, the unstandardized coefficients were standardized, then combined using a random-effect meta-analysis using Field and Gillett's (2010) SPSS macro, using the Hedges' et al. algorithm. Using this meta-analytic procedure, we could determine whether there are significant gender differences overall across all 14 analyses. If we successfully replicate Eastwick and Finkel's (2008) findings, we expect this gender difference calculated using meta-analysis to be non-significant.

Differences From Original Study

Due to external limitations, there were some deviations between the original study and our replication, which we note here: (a) We recruited participants with an incentive of extra credit in psychology courses; we could not offer monetary compensation as in the original design. (b) We were constrained by technological capacities of Internet servers at Maryland, and therefore, instead of using a customized

website for storing survey data and communication between participants, we used individualized Qualtrics surveys for pre- and post-event follow-ups. We also instructed participants to create anonymous email accounts for communication with potential partners in the study. (c) Participants were allowed to sign up for events up until 24 hr in advance of the event, thus, they completed the pre-event survey closer in time to the speed-dating event (1 to 4 days before the event), rather than 6 to 13 days before the event as in the original study. (d) The speed-dating events were held in two spacious auxiliary rooms on the Maryland campus (ordinarily used for social events), with comfortable seating, lighting, and music, rather than an art gallery as in the original study (which was not available to us). (e) One of the relationship initiation outcomes in the follow-up surveys (“I am eager to get to know [name] better”) was dichotomous (“yes/no”) in our study, but continuous on a 1 to 9 scale in the original study. (f) Due to human error, one of the four items assessing the degree to which participants rated their ideal partners and dates/matches as “personable” was missing from our follow-up surveys. We do not believe that any of these differences are substantial.

Results

Descriptive Statistics

Means, standard deviations, skewness, kurtosis, and ranges are presented in Table 1. Generally speaking, most variables were normally distributed and had sufficient variance to analyze. Notably, earning prospects and personable variables had a slight negative skew suggesting there were comparatively fewer people at the low end of these scales. In contrast, physical attractiveness was roughly normally distributed. Overall, there was sufficient variability to analyze; however, due to the negative skew, readers should be more cautious about inferences about very low levels of earning prospects and personable variables.

Intraclass correlations (Table 2) were used to demonstrate the percentage of variance available to explain at each respective level of the multilevel model. Because analyses take place primarily at the within-subjects level (Level 1), there must be significant variance to explain at Level 1 to proceed with analyses. Approximately 41% to 86% of the variance in the two-level models (interaction record and post-match) was at the within-subjects level (i.e., with each rating of a speed-date/match as an observation). In comparison, 23% to 60% of the variance in the three-level models (follow-ups) was at Level 1 (i.e., with each repeated-measure follow-up as an observation). Overall, the intraclass correlations supported analyzing data at the within-subjects level, although it must be acknowledged that significant between-subjects variance also existed (i.e., individual people had a tendency to rate all their speed-dates/matches in a similar way).

Table 1. Descriptive Statistics.

	<i>M</i>	<i>SD</i>	Skewness	Kurtosis	Observed range	<i>n</i>
Interaction record						
Earning prospects	6.33	1.82	-0.71	0.27	1-9	2,682
Physically attractive	4.61	2.25	-0.03	-1.05	1-9	2,702
Personable	6.45	1.51	-0.89	1.00	1-9	2,684
Romantic desire	5.13	1.89	-0.27	-0.67	1-9	2,763
Chemistry	4.92	1.80	-0.26	-0.39	1-9	2,761
“yessing”	0.28	0.45	—	—	—	2,629
Post-match						
Excitement	4.88	1.32	-0.63	0.80	1-7	242
Initiation plans	4.18	1.57	-0.14	-0.41	1-7	243
Initiation hopes	4.77	1.60	-0.59	-0.07	1-7	243
Follow-up (all matches)						
Earning prospects	6.06	1.63	-0.60	0.90	1-9	1,133
Physically attractive	5.33	1.87	-0.40	-0.05	1-9	1,133
Personable	6.44	1.68	-1.02	1.45	1-9	1,135
Get to know better	0.43	0.50	—	—	—	1,149
Follow-up (only if hangout)						
Date initiation	3.76	1.95	-0.06	-1.10	1-7	242
Date enjoyment	4.68	1.74	-0.62	-0.20	1-7	242
Follow-up (romantic only)						
Passion	4.09	1.11	0.22	-0.07	2-7	170
Commitment	4.16	1.65	0.07	-0.42	1-7	165
Desire for one-night stand	4.25	2.01	-0.10	-1.05	1-7	165
Desire for casual	3.73	1.45	0.30	-0.26	1-7	164
relationship						
Desire for serious	4.15	1.74	0.12	-0.85	1-7	165
relationship						
Follow-up (sexual)						
Sexual initiation	3.54	1.63	0.04	-0.16	1-7	78
Sexual enjoyment	4.76	1.87	-0.65	-0.33	1-7	75
Physical contact good idea	3.50	1.84	0.26	-0.65	1-7	78

Note. Partner characteristics were measured on a 1 to 9 scale with higher numbers indicating greater presence of the characteristic in the partner. Relationship initiation dependent variables were measured on a 1 to 7 scale, except for romantic desire and chemistry, which were measured on a 1 to 9 scale, and both “yessing” and “get to know better,” which were coded 1 for yes and 0 for no.

Internal consistency (Table 2) of multi-item scales was assessed using a multilevel approach to Cronbach’s alpha (Geldhof, Preacher, & Zyphur, 2014) in the two-level models. Overall, reliability of these scales was adequate to excellent (α s from .80 to .96). Because there is not currently any agreed-upon method for calculating internal consistency in a three-level data set, data were first averaged across all participants, then Cronbach’s alphas were calculated at Level 3. These alpha reliabilities were also excellent (α s from .85 to .96).

We also calculated the number of participants, matches, and follow-up surveys completed for each subsample of participants. Because of the complexities of the design, there were six subsamples, each contingent on various pivot questions. The first subsample is the “Interaction Record,” which includes participants’ ratings of each of their speed dates before the matching process ($N = 307$) with an average of 9.15 speed dates.⁴ The second subsample is the “Post-Match,”

which includes only participants who mutually said “yes” to further correspondence with each other. Around 43.0% of participants ($n = 132$) completed the post-match questionnaire, and participants had 1.84 matches on average. Overall, this is somewhat lower than the number of matches found in Eastwick and Finkel’s (2008) original article ($M = 2.54$ matches).

In the third subsample, “Follow-up all matches,” 46.9% of the participants ($n = 144$) had found a match,⁵ with 1.72 matches on average, and participants completed an average of 4.67 of a maximum of 10 follow-up surveys. In the fourth subsample, “Follow-up only if hangout,” only participants who hung out or corresponded with the match completed questions ($n = 58$; average number of matches corresponded with = 1.33; number of usable follow-up surveys per match = 3.16). In the fifth subsample, “Follow-up romantic,” only participants who rated the match as a romantic partner, or a person with romantic potential filled

Table 2. Intraclass Correlations and Alpha Reliabilities.

	No. of items	ICC _{lev3}	ICC _{lev2}	ICC _{lev1}	α_{lev3}	α_{lev2}	α_{lev1}
Interaction record							
Earning prospects	2	—	.28	.72	—	.96	.80
Physically attractive	2	—	.38	.62	—	.92	.89
Personable	4	—	.37	.63	—	.95	.84
Romantic desire	3	—	.31	.69	—	.80	.86
Chemistry	3	—	.27	.73	—	.95	.90
“yessing”	1	—	.14	.86	—	—	—
Post-match							
Excitement	1	—	.49	.51	—	—	—
Initiation plans	1	—	.56	.44	—	—	—
Initiation hopes	1	—	.59	.41	—	—	—
Follow-up (all matches)							
Earning prospects	2	.39	.24	.37	.96	—	—
Physically attractive	2	.43	.34	.23	.92	—	—
Personable	3	.41	.14	.45	.96	—	—
Get to know better	1	.31	.27	.42	—	—	—
Follow-up (only if hangout)							
Date initiation	1	.57	.13	.30	—	—	—
Date enjoyment	1	.02	.58	.40	—	—	—
Follow-up (romantic only)							
Passion	5	.57	.02	.41	.85	—	—
Commitment	1	.38	.08	.54	—	—	—
Desire for one-night stand	1	.39	.01	.60	—	—	—
Desire for casual relationship	1	.31	.13	.56	—	—	—
Desire for serious relationship	1	.54	.01	.45	—	—	—
Follow-up (sexual)							
Sexual initiation	1	.58	.09	.40	—	—	—
Sexual enjoyment	1	.77	.02	.21	—	—	—
Physical contact good idea	1	.62	.02	.36	—	—	—

Note. Each bolded heading begins a separate subgroup of participants (see “Results” section for more details). ICC_{lev3} indicates the percent variance that exists to explain at Level 3. ICC_{lev2} indicates the percent variance that exists to explain at Level 2. ICC_{lev1} indicates the percent variance that exists to explain at Level 1. α_{lev3} , α_{lev2} , and α_{lev1} indicate Cronbach’s alpha (internal consistency) calculated for each multi-item scale at each level of the data. As there is not currently any agreed-upon method for calculating reliability in a three-level data set, data were first averaged across all participants, then alphas were calculated to get α_{lev3} . Values of “—” indicate that a particular statistic cannot be calculated for this particular variable due to the multilevel structure of the data set and/or complexities of the design. ICC = intraclass correlation.

out questions ($n = 41$; average number of matches in romantic relationship = 1.17, average number of usable follow-up surveys per match = 3.45). In the sixth and final subsample, “Follow-up sexual,” only participants who had sexual contact with their partner filled out the questionnaire ($n = 13$; average number of matches with sexual contact = 1.08; average number of sexual follow-up surveys completed = 2.64). As in Eastwick and Finkel’s (2008) article, too few participants had sexual contact with matches, so these questions were not analyzed further. Moreover, there were too few write-ins ($n = 34$) to analyze separately; Eastwick and Finkel’s (2008) original article had 143 write-ins. However, all other subsamples were analyzed.

Earning Prospects Analyses

Table 3 contains the confirmatory analyses testing whether gender moderates the relationship between earning prospects

and 14 relationship initiation variables. Overall, 9 of 14 relationships for men were statistically significant and positive. This is broadly consistent with Eastwick and Finkel’s (2008) findings that 9 of 14 relationships were positive and statistically significant. In contrast, 8 of 14 analyses for women were statistically significant. Eastwick and Finkel (2008) found 11 of 14 analyses for women were positive and statistically significant. Only two gender by earning prospects interaction effects were statistically significant in the present data (i.e., when predicting date initiation and “get to know better”; in both cases, the effect was larger for men than for women). In contrast, Eastwick and Finkel (2008) found four significant interactions suggesting that effect sizes are larger for men, and one significant interaction suggesting effect sizes are larger for women.

Because some individual analyses had relatively small sample sizes, the overall meta-analytic summary is a better test of hypotheses, as it weights each finding according to

Table 3. Gender Differences in the Relationship Between Earning Prospects and Relationship Initiation Variables.

Relationship initiation variables	Male B (SE)	Female B (SE)	Male <i>r</i>	Female <i>r</i>	Male <i>n</i> _{obs}	Female <i>n</i> _{obs}
Interaction record						
Romantic desire	0.31 (0.04)	0.31 (0.03)	.32***	.37***	1,334	1,344
Chemistry	0.31 (0.04)	0.33 (0.03)	.34***	.38***	1,334	1,334
“yessing”	0.30 (0.05)	0.23 (0.06)	.08***	.06***	1,260	1,246
Post-match						
Excitement	0.18 (0.06)	0.19 (0.06)	.34**	.30**	114	116
Initiation plans	0.19 (0.08)	0.22 (0.07)	.29*	.37**	114	117
Initiation hopes	0.06 (0.09)	0.24 (0.07)	.11	.35**	114	117
Follow-up (all)						
Get to know better	0.41 (0.04)	0.34 (0.09)	.11***	.09***	514	614
Follow-up (hangout)						
Date initiation	0.45 (0.10)	-0.09 (0.17)	.54***	-.18	111	124
Date enjoyment	0.41 (0.14)	-0.03 (0.32)	.63**	-.05	110	125
Follow-up (romantic)						
Passion	-0.06 (.12)	0.16 (0.17)	-.11	.38*	86	79
Commitment	-0.16 (0.19)	0.11 (0.09)	-.21	.14	86	79
Desire for one-night stand	0.33 (0.20)	-0.06 (2.43)	.40*	-.06	86	79
Desire for casual relationship	0.14 (0.15)	0.00 (0.13)	.21	.00	86	78
Desire for serious relationship	-0.15 (0.17)	-0.04 (0.11)	-.19	-.07	86	79
Overall <i>r</i> [95% CI]			.22 [.11, .33]	.16 [.04, .28]		
Overall <i>r</i> (both sexes combined)			.19 [.12, .26]			
Overall significance of interaction			$\chi^2(1) = 0.66, p = .42$			

Note. Regression Bs indicate the relationship earning prospects and 14 relationship initiation dependent variables, which were regressed in separate analyses. Partner characteristics were measured on a 1 to 9 scale with higher numbers indicating greater presence of the characteristic in the partner. Relationship initiation dependent variables were measured on a 1 to 7 scale, except for romantic desire and chemistry, which were measured on a 1 to 9 scale, and both “yessing” and “get to know better,” which were coded 1 for yes and 0 for no. Overall *r*s were calculated using Hedges’ random-effect meta-analysis. Significant interaction effects are indicated by bolded text (i.e., is gender a moderating variable?). The overall significance of gender as a moderator was calculated using meta-analysis with continuous variables as random effects and gender as a fixed effect. *n*_{obs} indicates the number of observations used within analyses, and were used for weighting in the meta-analysis. Analyses for write-ins and sexual variables omitted due to small sample sizes. CI = confidence interval.

p* < .05. *p* < .01. ****p* < .001.

the number of observations and reduces noise by averaging across all analyses. Overall, the meta-analytic summary of these results showed that perceived earning prospects significantly predicted relationship initiation variables with a very similar effect size to Eastwick and Finkel (2008), with *r* = .18 in the original article, and *r* = .19 in the present replication. This relationship was not moderated by gender of the participant, although the data trended slightly toward the tendency for the effect sizes to be stronger for men when compared with women, contrary to predictions made by evolutionary models, but consistent with Eastwick and Finkel’s (2008) original findings. Overall, the meta-analytic results suggest that the present replication was successful.⁶

Physical Attractiveness Analyses

Table 4 contains additional analyses testing whether gender moderates the relationship between physical attractiveness and 14 relationship initiation variables. Overall, 9 of 14 relationships for both men and women were statistically significant and positive, with 1 significant negative relationship for women. In contrast, Eastwick and Finkel (2008) found that 12

of 14 relationships were positive and statistically significant for men, and 13 of 14 were positive and significant for women.

None of the interaction effects was statistically significant in the present data. Eastwick and Finkel’s (2008) found two significant interactions suggesting that effect sizes are larger for men, and three significant interaction suggesting effect sizes are larger for women. Again, the overall meta-analytic summary is a better test of hypotheses. The meta-analysis was conducted in the same fashion as for earning prospects. Overall, the meta-analytic summary of these results showed that perceived physical attractiveness significantly predicted relationship initiation variables with a very similar effect size to Eastwick and Finkel (2008), with *r* = .43 in the original article, and *r* = .36 in the present replication. As with earning prospects, this relationship was not moderated by gender of the participant, $\chi^2(1) = 0.07, p = .80$. Thus, the meta-analytic results suggest that these analyses also successfully replicated Eastwick and Finkel (2008).

Personable Analyses

Table 5 contains additional analyses testing whether gender moderates the relationship between “personable” and 14

Table 4. Gender Differences in the Relationship Between Physical Attractiveness and Relationship Initiation Variables.

Relationship initiation variables	Male B (SE)	Female B (SE)	Male <i>r</i>	Female <i>r</i>	Male <i>n</i> _{obs}	Female <i>n</i> _{obs}
Interaction record						
Romantic desire	0.60 (0.40)	0.62 (0.03)	.73***	.73***	1,353	1,345
Chemistry	0.47 (0.03)	0.52 (0.02)	.63***	.65***	1,353	1,345
“yessing”	0.80 (0.08)	0.67 (0.07)	.21***	.18***	1,268	1,257
Post-match						
Excitement	0.27 (0.07)	0.26 (0.05)	.54***	.51***	115	115
Initiation plans	0.26 (0.08)	0.18 (0.07)	.43***	.36*	115	116
Initiation hopes	0.23 (0.08)	0.33 (0.07)	.45***	.56***	115	116
Follow-up (all)						
Get to know better	0.30 (0.09)	0.29 (0.09)	.08***	.08***	471	601
Follow-up (hangout)						
Date initiation	0.42 (0.17)	0.27 (0.07)	.52**	.57***	111	124
Date enjoyment	0.51 (0.09)	0.54 (0.09)	.76***	.80***	110	125
Follow-up (romantic)						
Passion	0.07 (0.15)	-0.15 (0.08)	.09	-.35*	86	79
Commitment	0.04 (0.15)	-0.09 (0.10)	.04	-.12	86	79
Desire for one-night stand	0.08 (0.10)	-0.07 (0.16)	.08	-.06	86	79
Desire for casual relationship	0.06 (0.15)	-0.07 (0.08)	.07	-.10	86	78
Desire for serious relationship	-0.06 (0.23)	0.05 (0.12)	-.06	.07	86	79
Overall <i>r</i> [95% CI]			.38 [.16, .61]	.34 [.10, .59]		
Overall <i>r</i> (both sexes combined)			.36 [.21, .52]			
Overall significance of interaction			$\chi^2(1) = 0.07, p = .80$			

Note. Regression Bs indicate the relationship earning prospects and 14 relationship initiation dependent variables, which were regressed in separate analyses. Partner characteristics were measured on a 1 to 9 scale with higher numbers indicating greater presence of the characteristic in the partner. Relationship initiation dependent variables were measured on a 1 to 7 scale, except for romantic desire and chemistry, which were measured on a 1 to 9 scale, and both “yessing” and “get to know better,” which were coded 1 for yes and 0 for no. Overall *r*s were calculated using Hedges’ random-effect meta-analysis. There were no significant interactions in Table 4 (i.e., gender was not a moderating variable). The overall significance of gender as a moderator was calculated using meta-analysis with continuous variables as random effects and gender as a fixed effect. *n*_{obs} indicates the number of observations used within analyses, and was used for weighting in the meta-analysis. Analyses for write-ins and sexual variables omitted due to small sample sizes. CI = confidence interval.

p* < .05. *p* < .01. ****p* < .001.

relationship initiation variables. Overall, 7 of 14 relationships for men and 9 out of 14 relationships for women were statistically significant and positive, with 1 significant negative relationship for women. Eastwick and Finkel (2008) found that 10 of 14 relationships were positive and statistically significant for men, and 13 of 14 were positive and significant for women.

Only two gender by personable interaction effects were statistically significant in the present data (i.e., when predicting desire for a casual relationship and “get to know better”; effect sizes were more positive for women than for men). Eastwick and Finkel (2008) found three significant interactions (chemistry, initiation plans, initiation hopes), also suggesting that effect sizes are larger for women. The meta-analysis was conducted in the same fashion as for earning prospects and attractiveness in Tables 2 and 3. Overall, results showed that perceived personable characteristics significantly predicted relationship initiation variables with a similar effect size to Eastwick and Finkel (2008), with *r* = .26 in the original article, and *r* = .29 in the present replication. As with earning prospects and physical attractiveness, this relationship was not moderated by gender of the participant,

$\chi^2(1) = 0.06, p = .80$. Thus, the meta-analytic results suggest that these analyses also successfully replicated Eastwick and Finkel (2008).

Summary of Replication Attempt

In this study, we fully replicated the primary findings from the original (Eastwick & Finkel, 2008) study. Specifically, we found that the perception of greater earning prospects, physical attractiveness, and warmth in potential partners was associated with greater romantic interest in those partners (*r* = .19 [.12, .26]; *r* = .36 [.21, .52]; and *r* = .29 [.15, .44], respectively). However, none of these effects was moderated by gender, $\chi^2(1) = 0.66, p = .42$; $\chi^2(1) = 0.07, p = .80$; and $\chi^2(1) = 0.06, p = .80$, respectively. In this case, the key results (which we replicated) were three non-significant two-way interactions between perceived (X1) partner earning prospects, (X2) physical attractiveness, (X3) warmth, and (Y) romantic interest, with gender as the moderator. The alternative hypothesis was that gender would significantly moderate the bivariate associations in the manner espoused by evolutionary psychologists (e.g., Buss & Schmitt, 1993;

Table 5. Gender Differences in the Relationship Between Being Personable and Relationship Initiation Variables.

Relationship initiation variables	Male B (SE)	Female B (SE)	Male <i>r</i>	Female <i>r</i>	Male <i>n</i> _{obs}	Female <i>n</i> _{obs}
Interaction record						
Romantic desire	0.72 (0.04)	0.73 (0.03)	.57***	.63***	1,341	1,339
Chemistry	0.77 (0.04)	0.77 (0.03)	.66***	.72***	1,341	1,339
“yessing”	0.73 (0.09)	0.69 (0.09)	.20***	.19***	1,257	1,251
Post-match						
Excitement	0.24 (0.09)	0.25 (0.10)	.33**	.32**	113	117
Initiation plans	0.22 (0.12)	0.38 (0.12)	.25	.48**	113	118
Initiation hopes	0.09 (0.11)	0.40 (0.12)	.12	.47**	113	118
Follow-up (all)						
Get to know better	0.57 (0.05)	0.54 (0.07)	.16***	.15***	513	617
Follow-up (hangout)						
Date initiation	0.21 (0.15)	0.39 (0.13)	.29	.47***	110	124
Date enjoyment	0.47 (0.15)	0.72 (0.19)	.70***	.66***	110	125
Follow-up (romantic)						
Passion	-0.06 (0.12)	-0.29 (0.10)	-.10	-.34**	86	79
Commitment	-0.16 (0.19)	0.11 (0.09)	-.21	.14	86	79
Desire for one-night stand	0.33 (0.20)	-0.06 (2.43)	.40*	-.06	86	79
Desire for casual relationship	0.12 (0.13)	-0.16 (0.17)	.18	-.12	86	78
Desire for serious relationship	-0.13 (0.18)	0.18 (0.18)	-.17	.13	86	79
Overall <i>r</i> [95% CI]			.28 [.08, .47]	.31 [.07, .55]		
Overall <i>r</i> (both sexes combined)			.29 [.15, .44]			
Overall significance of interaction			$\chi^2(1) = 0.06, p = .80$			

Note. Regression Bs indicate the relationship earning prospects and 14 relationship initiation dependent variables, which were regressed in separate analyses. Partner characteristics were measured on a 1 to 9 scale with higher numbers indicating greater presence of the characteristic in the partner. Relationship initiation dependent variables were measured on a 1 to 7 scale, except for romantic desire and chemistry, which were measured on a 1 to 9 scale, and both “yessing” and “get to know better,” which were coded 1 for yes and 0 for no. Overall *r*s were calculated using Hedges’ random-effect meta-analysis. Significant interaction effects are indicated by bolded text (i.e., is gender a moderating variable?). The overall significance of gender as a moderator was calculated using meta-analysis with continuous variables as random effects and gender as a fixed effect. *n*_{obs} indicates the number of observations used within analyses, and was used for weighting in the meta-analysis. Analyses for write-ins and sexual variables omitted due to small sample sizes. CI = confidence interval.

p* < .05. *p* < .01. ****p* < .001.

Fisman et al., 2006), but we did not find support for this hypothesis. In addition, the aggregate effect size for the main effects (e.g., the association between earning prospects and romantic interest) was comparable with the effect size reported in the original study. This strengthens our confidence in the true effect reported by the original study authors (*r* = .19 and *r* = .16 for men and women, respectively), as well as more recent meta-analyses (Eastwick et al., 2014).

Discussion

Our study adds further evidence to support the claim that stated ideal partner preferences, although robust and relevant to evolutionary psychological literature, have less predictive value when considering attraction and dating in a live context. Previous work has shown that in the heterosexual population, men and women report mate preferences (for physical attractiveness and earning prospects) that are statistically different from each other, but when men and women meet face to face, the associations between partner traits and romantic interest are statistically equivalent across gender. This is consistent with more recent theory about the nature of romantic

attraction (Eastwick et al., 2014). Specifically, construal-level theory (Trope & Liberman, 2003, 2010) predicts that people tend to evaluate specific stimuli (in this case, when meeting potential romantic partners in person) differently than they would conceptualize abstract ideas (in this case, general preferences for what they would prefer in an ideal romantic partner), thus accounting for why divergent results exist when examining hypothetical and actual dating experiences.

Furthermore, although explicit preferences tend to be poorer predictors of face-to-face romantic attraction, implicit “gut-level” preferences (measured by computer-based reaction time tasks) provide better predictors of face-to-face romantic attraction. This does not mean explicit preferences are completely without predictive value. In fact, Eastwick, Eagly, Finkel, and Johnson (2011) postulated that explicit preferences may play a larger role in predicting other dating/relationship outcomes, aside from initial feelings of romantic attraction. As a recent example (Eastwick & Neff, 2012), explicit reports of ideal partner preferences significantly predict divorce rates when examined as *patterns* (e.g., preferring warmth over ambition in potential partners) but not as

levels (e.g., preferring high amounts of warmth or ambition in potential partners).

As we noted above, the existing literature utilizing live-attraction methodologies has yielded mixed results, with some studies yielding significant interactions between participant gender, partner characteristics, and attraction/desire in the direction predicted by evolutionary theory (e.g., Li et al., 2013), whereas others (e.g., Eastwick & Finkel, 2008, and the present replication) have failed to find such supporting evidence. However, studies that have examined long-term relationship outcomes have found support for the idea that spousal partner selection is driven by sex-specific preferences consistent with parental investment theory. For example, Meltzer, McNulty, Jackson, and Karney (2014b) found that for men, the physical attractiveness of their spouses significantly predicted their marital satisfaction over a 4-year period to a much greater extent compared with women. However, the authors themselves note that it is not clear whether physical attractiveness is always more important to men than women; it may be that physical attractiveness matters equally to men and women in short-term mating contexts, but that physical attractiveness matters more to men in long-term mating contexts (Meltzer, McNulty, Jackson, & Karney, 2014a). Thus, direct evidence to date supporting the tenets of evolutionary theories for heterosexual romantic attraction is mixed. A recent meta-analysis (Eastwick et al., 2014) summarizing data from 97 studies (tens of thousands of data points) and a variety of different methodologies does not find empirical support for the claim that heterosexual men's and women's experiences of romantic attraction reliably differ according to potential partners' physical attractiveness or earning prospects. When placing our findings into proper context with prior research, we place more weight on meta-analytic results. Thus, although we have reviewed some evidence in favor of sex differences in mate preferences, we believe that the bulk of the evidence (including the present study's results) suggests that there are not sex differences in mate preferences in face-to-face, short-term mating contexts. Indeed, it appears as though men and women both tend to prefer a partner who is attractive and personable, and has high earning prospects.

Limitations and Caveats

There are a few statistical caveats to the current study worth mentioning, to put the key results into proper context. First, the key results that we replicated from Eastwick and Finkel (2008) were derived through meta-analysis, aggregating across a variety of multilevel analyses. Some of the specific two-way interactions for unique predictor variables were significant as reported in the original article, but none of those specific interactions replicated in our sample. This suggests that the individual interactions here and in Eastwick and Finkel (2008) are likely false positives. Much of the variability across analyses is stabilized when they are combined

through meta-analysis, and the overall estimate of effect size calculated this way is very consistent with the original article's results. This pattern of results highlights the importance of a pre-registered analysis plan, and demonstrates how incorrect conclusions could be made by exploiting researcher degrees of freedom by hand-picking specific dependent variables (DVs) post hoc from a wide variety of measured outcomes. For instance, if we were to present only the "Get to know better" and "date initiation" analyses for earning prospects, we may have (incorrectly) concluded that men value earning prospects more than women.

Overall, our sample ($N = 307$) was adequately powered to detect the three-way meta-analytic interaction (if it existed), given the assumptions laid out in the "Method" section. However, not all participants received matches, and not all participants who received matches actually followed up with them. As reported in Table 2, the data set shrunk progressively for each category of follow-up variables (e.g., "hang-out" $n = 81$; "romantic" $n = 55$; "sexual" $n = 26$). Thus, fewer and fewer participants were able to contribute data for these specific analyses. According to our power estimates, we would need roughly 687 data points per analysis (i.e., 275 participants \times 2.5 matches calculated in the power analyses) to detect a significant effect (if one existed) for that variable, and most of our analyses fell substantially short of that number.

There is an important theoretical distinction between general romantic interest (as an abstract construct) and the specific subtype behaviors that fall within that category (e.g., "desire for a one-night stand"), especially considering the theoretical propositions made by Eastwick et al. (2014) stemming from construal-level theory (Trope & Liberman, 2003, 2010). Our conclusion focuses on general romantic interest and not the specific subtype behaviors. Theoretically, one or more of these specific effects could exist in the world, despite a null effect for the general romantic interest construct, but we would have no way of knowing based on our underpowered data (the likelihood of a Type 2 error is substantial). As an example, perhaps there exists a greater association between perceived earning prospects and the likelihood of sexual enjoyment for women compared with men. We are too underpowered to draw firm conclusions about these more specific (non-significant) interaction effects. It should be noted (see Table 3) that we did not find evidence for a two-way interaction in any of the specific relationship initiation variables assessed through the interaction records (e.g., "Romantic desire") for which we were adequately powered. Future studies could investigate (with very large samples) potential sex differences in explicit ideal mate preferences and short-term sexual outcomes using similar methodologies. Again, we must strongly emphasize that this statistical power caveat does not call into question the original results reported by Eastwick and Finkel (2008) because their key result was derived from meta-analytic data, which is an effective way of summarizing numerous effect

sizes, even if individual tests are underpowered (Field & Gillett, 2010). We have not concluded that our data were inadequate to address the replication, but merely that we wish to call attention to a subtle nuance in the method that precludes us from making definitive statements about all subtypes of specific romantic interest outcomes (e.g., sexual contact). Our conclusion is based on general romantic interest as an aggregated construct composed of many outcomes.

Although our replication of Eastwick and Finkel's (2008) study did reproduce their key results, the null interaction effect we found may have had more to do with lower return rate of surveys, as well as the lack of mutual matches, and (to a lesser extent) participants' low intrinsic motivation to participate in a speed-dating experiment. Anecdotally, some of our participants remarked that they would not ordinarily choose to participate in a speed-dating event (due to social stigma), and many agreed to participate only to receive compensation (extra credit) or to support their friends. Some participants were heard to remark that they were not interested in dating; one stated that she was in a relationship and not genuinely interested in new partners, while another stated he was homosexual. All participants in our study explicitly identified as heterosexual and single on survey materials, so we do not have an accurate estimate for how endemic this was in our sample, nor do we have grounds to exclude any participant's data. Many other participants did have genuinely positive dating experiences, and some reported that they were still dating one of the people they met during an event, months after the study concluded. So ultimately, mixed anecdotal reports from student participants do not undermine the general conclusions drawn from this sample. We did not survey our participants in their attitudes toward speed dating or intrinsic motivation to speed date, thus precluding our ability to draw firm empirical conclusions. We strongly suggest future studies investigate norms and possible stigmas associated with speed dating in undergraduate populations, as they may be dissimilar from the dating norms in older adults, thus potentially decreasing the generalizability of these results (and other studies).

The methodological information provided by this study may be particularly useful to researchers who aspire to use speed-dating methodologies in the future. If our results are suggestive of speed dating more typically, researchers who are interested in examining sexual initiation as an outcome may need to collect extremely large samples to achieve adequate power when using a speed-dating paradigm because relatively few speed-dating participants actually formed a romantic relationship with one of their dates. This may reflect a general weakness in the research methodology, or it may reflect a phenomenon that is endemic to the young adult population or college student atmosphere, where casual relationships and hookups at parties are more the norm, whereas organized dating events are not.

There are many people who are eager and excited to participate in speed-dating events outside the college/university

environment, and will even pay the event organizers for the opportunity to meet and date new people. This is the exact opposite of the model used in past research (Eastwick & Finkel, 2008; Finkel & Eastwick, 2008) and in the present study, where we compensated participants for coming to events. We suggest that the college student population is less adequate for speed-dating research compared with community samples, due to college culture norms and peer-mediated stigmas associated with speed dating. We also suggest that researchers interested in utilizing this method should attempt to form partnerships with existing organizations that conduct dating events (e.g., groups on Meetup.com), which would yield data that are more externally valid compared with data collected from undergraduate populations in which participants are less interested in formal dating.

Regardless of participants' stated interest in a long-term or short-term relationship, they may perceive dating partners as suitable more for short-term relationship potential if they meet in a speed-dating context, and thus, the sex-specific preferences predicted by evolutionary theory/SST to explain long-term mate choice would not be visible (Kurzban & Weeden, 2007). It may also be the case that participants who participate in speed-dating events are not representative of the general dating population (Asendorpf, Penke, & Back, 2011), or they may not represent the full spectrum of population variability in traits such as physical attractiveness or potential earning prospects (especially for college students; see Li et al., 2013).

Some new lines of research have noted these methodological limitations and attempted to improve on them. For example, Li et al. (2013) used a modified speed-dating paradigm with an experimental manipulation of perceived social status and physical attractiveness, and found sex differences consistent with evolutionary theories of parental investment. They argued that this pattern of results will only emerge if participants include the full spectrum of variability found in the general population and participants are asked to rate others using the general population as a reference group. They contend that, even if there appears to be substantial variability in responding in the descriptive statistics, a lack of participants on the low end of traits will be masked if participants do not rate others relative to the general population. For example, they show in Study 1 that low earning prospect students are rated *average* when compared with the general population. Thus, the present study as well as numerous other prior studies on this topic might be criticized on these grounds.

Cultural differences might also moderate the degree to which men's and women's attraction to others reflects their stated preferences. Consistent with this idea, other research has found that differences in mate preferences for men and women shrink as a function of their nation's economic parity (Zentner & Mitura, 2012). However, the impact of those socio-cultural variables on attraction outcomes based on gender does not necessarily challenge or weaken evolutionary

explanations. One possibility is that evolutionary processes designed facultative mechanisms that adjust preferences based on exposure and availability of different types of mates. In this model, individuals' preferences for traits in potential partners would develop somewhat differently in places where mates are particularly high/low on attractiveness, earning prospects, and so on, and may be dependent on local conditions or opportunities (e.g., women's preferences for resources in men may adaptively decrease if the women themselves have direct access to resources). Thus, an evolutionary explanation is not precluded. Culture environment and evolution are intertwined and not necessarily mutually exclusive.

Conclusion

Ultimately, lack of supporting data is not evidence for any research hypothesis, and none of the methodological/statistical caveats listed above undermines the theoretical propositions and conclusions from the original study. Eastwick and Finkel's (2008) research was meant to address an alternative theory regarding gender and evolutionary-derived *sexual strategies* in human mating. Importantly, we did not find any support for that theory in our data, consistent with recent meta-analytic data on this topic (Eastwick et al., 2014). If anything, a trend in our data suggests the opposite of what SST (Buss & Schmitt, 1993) and other studies (e.g., Fisman et al., 2006) would propose; the association between earning prospects and romantic interest was larger for men than for women. Although in contrast, the association between physical attractiveness and romantic interest was also larger for men than for women, which is consistent with SST. However, given that these gender differences were non-significant (and that the confidence intervals overlapped substantially), they should not be interpreted with confidence. Overall, we found no reliable indication from our data to support the predictions laid out by SST. Our study adds to the "replicability" of these null effects, provides grounds for methodological innovations for future speed-dating research, and adds to the rich literature on romantic attraction and dating processes.

Acknowledgment

Mary Furda, Jenn Loya, Ed Walsh, Delaram Mirzania, Jessica Maclin, Michael Shrager, Kapil Chauhan, Rayna Carranza, Jake Moore, Chelsea Brown, Christin Probst, Katherine Cook, Rabee Haq, Lisa Patti, Tara Schluskel, Erin DeVita, Shani Banks, Jessica Maxin, Lisa MacHarrie, Brooke Rothman, Nicole Travers, James Riffle, Sphoorti Pandit, Nina Shen, and Sam Kravitz are thanked for their research assistance.

Authors' Note

We have been in contact with the original study authors consistently from the inception of this replication attempt, and we are pleased to report their warm, accommodating demeanor, and full-hearted

cooperation with our research. They provided us with raw materials, syntax, methodological advice, and positive reinforcement. Eastwick and Finkel should be commended as good scientists.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: This study was funded in part by the Pride Alliance at the University of Maryland. Sean P. Mackinnon was supported by a postdoctoral fellowship from the Social Sciences and Humanities Research Council (SSHRC).

Notes

1. Note that this power analysis is for a two-level model. We also conduct some analyses with three levels, by adding repeated measures. However, because the addition of repeated measures will increase the statistical power of analyses, we based our sample size estimates on the least powerful analysis type in our plan (i.e., the two-level analyses) to remain conservative.
2. The language used for the write-in option was as follows: "You now have the opportunity to add, or 'write-in,' any individuals who you met outside of the speed-dating event and are currently interested in. If you would like to add a 'write in' (aka someone from outside the study), Select 'write in.'"
3. Note that "yessing" is a dichotomous variable. Thus, the analysis using this variable requires the use of logistic regression. When standardizing the B values in meta-analysis at a later step, we will follow Eli and Finkel's original procedure for standardizing the B values: "the B estimate . . . was first converted to a d ($B \times \sqrt{3/\pi}$) and then to an r ($\sqrt{d^2 / (d^2 + 4)}$), and finally the variance of this correlation was calculated with the formula $(1 - r^2)^2 / (N - 2)$ (Haddock, Rindskopf, & Shadish, 1998; Rosenthal, 1994)" (Eastwick & Finkel, 2008, p. 252).
4. The sample sizes and number of matches here are representative of the data actually analyzed, and because of missing data, will not match the sample size and number of matches in the participants section exactly.
5. There are more matches at follow-up because this subsample had the opportunity to contact "missed matches" (i.e., where only one person said yes initially, but the other person later changed his or her mind).
6. To ensure that our results were not due to minor differences in statistical software, we also performed the meta-analysis in SAS® using the original syntax provided by the original authors. These results showed virtually identical effect size results for men ($r = .24$) and women ($r = .21$), and the interaction remained non-significant, $t(26) = 0.19$, $p = .85$. In the present article, we preferred to report the results from the Field and Gillett (2010) SPSS macro because this approach

also provides confidence intervals for main effects. However, it should be noted that the effect size from this supplementary analysis in SAS was ultimately the one chosen to include in the broader Reproducibility Project analyses, because it was deemed to be most directly comparable with the effect size from the original article, which also used SAS.

References

- Asendorpf, J. B., Penke, L., & Back, M. D. (2011). From dating to mating and relating: Predictors of initial and long-term outcomes of speed-dating in a community sample. *European Journal of Personality, 25*, 16-30. doi:10.1002/per.768
- Bosker, R. J., Snijders, T. A. B., & Guldmond, H. (2003). PINT (Power IN Two-level designs): *Estimating standard errors of regression coefficients in hierarchical linear models for power calculations* (user's manual, version 2.1). Retrieved from <http://stat.gamma.rug.nl/snijders/>
- Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review, 100*, 204-232. doi:10.1037/0033-295X.100.2.204
- Buss, D. M., Shackelford, T. K., Kirkpatrick, L. A., & Larsen, R. J. (2001). A half century of mate preferences: The cultural evolution of values. *Journal of Marriage and Family, 63*, 491-503. doi:10.1111/j.1741-3737.2001.00491.x
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Eastwick, P. W., Eagly, A. H., Finkel, E. J., & Johnson, S. E. (2011). Implicit and explicit preferences for physical attractiveness in a romantic partner: A double dissociation in predictive validity. *Journal of Personality and Social Psychology, 101*(5), 993-1011.
- Eastwick, P. W., & Finkel, E. J. (2008). Sex differences in mate preferences revisited: Do people know what they initially desire in a romantic partner? *Journal of Personality and Social Psychology, 94*, 245-264. doi:10.1037/0022-3514.94.2.245
- Eastwick, P. W., Luchies, L. B., Finkel, E. J., & Hunt, L. L. (2014). The predictive validity of ideal partner preferences: A review and meta-analysis. *Psychological Bulletin, 140*, 623-665. doi:10.1037/a0032432
- Eastwick, P. W., & Neff, L. A. (2012). Do ideal partner preferences predict divorce? A tale of two metrics. *Social Psychological & Personality Science, 3*, 667-674. doi:10.1177/1948550611435941
- Feingold, A. (1992). Gender differences in mate selection preferences: A test of the parental investment model. *Psychological Bulletin, 112*, 125-139. doi:10.1037/0033-2909.112.1.125
- Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology, 63*, 665-694. doi:10.1348/000711010X502733
- Finkel, E. J., & Eastwick, P. W. (2008). Speed-dating. *Current Directions in Psychological Science, 17*(3), 193-197.
- Finkel, E. J., Eastwick, P. W., & Matthews, J. (2007). Speed-dating as an invaluable tool for studying romantic attraction: A methodological primer. *Personal Relationships, 14*, 149-166. doi:10.1111/j.1475-6811.2006.00146.x
- Fisman, R., Iyengar, S. S., Kamenica, E., & Simonson, I. (2006). Gender differences in mate selection: Evidence from a speed dating experiment. *The Quarterly Journal of Economics, 121*, 673-697. doi:10.1162/qjec.2006.121.2.673
- Fletcher, G. O., Kerr, P. G., Li, N. P., & Valentine, K. A. (2014). Predicting romantic interest and decisions in the very early stages of mate selection: Standards, accuracy, and sex differences. *Personality and Social Psychology Bulletin, 40*, 540-550. doi:10.1177/0146167213519481
- Fletcher, G. O., Simpson, J. A., Thomas, G., & Giles, L. (1999). Ideals in intimate relationships. *Journal of Personality and Social Psychology, 76*, 72-89. doi:10.1037/0022-3514.76.1.72
- Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods, 19*, 72-91. doi:10.1037/a0032138
- Haddock, C. K., Rindskopf, D., & Shadish, W. R. (1998). Using odds ratios as effect sizes for meta-analysis of dichotomous data: A primer on methods and issues. *Psychological Methods, 3*, 339-353. doi:10.1037/1082-989X.3.3.339
- Kurzban, R., & Weeden, J. (2005). HurryDate: Mate preferences in action. *Evolution & Human Behavior, 26*, 227-244. doi:10.1016/j.evolhumbehav.2004.08.012
- Kurzban, R., & Weeden, J. (2007). Do advertised preferences predict the behavior of speed daters? *Personal Relationships, 14*, 623-632. doi:10.1111/j.1475-6811.2007.00175.x
- Li, N. P., Yong, J. C., Tov, W., Sng, O., Fletcher, G. O., Valentine, K. A., . . . Balliet, D. (2013). Mate preferences do predict attraction and choices in the early stages of mate selection. *Journal of Personality and Social Psychology, 105*, 757-776. doi:10.1037/a0033777
- Luo, S., & Zhang, G. (2009). What leads to romantic attraction: Similarity, reciprocity, security, or beauty? Evidence from a speed-dating study. *Journal of Personality, 77*, 933-964. doi:10.1111/j.1467-6494.2009.00570.x
- Meltzer, A. L., McNulty, J. K., Jackson, G. L., & Karney, B. R. (2014a). Men still value physical attractiveness in a long-term mate more than women: Rejoinder to Eastwick, Neff, Finkel, Luchies, and Hunt (2014). *Journal of Personality and Social Psychology, 106*, 435-440. doi:10.1037/a0035342
- Meltzer, A. L., McNulty, J. K., Jackson, G. L., & Karney, B. R. (2014b). Sex differences in the implications of partner physical attractiveness for the trajectory of marital satisfaction. *Journal of Personality and Social Psychology, 106*, 418-428. doi:10.1037/a0034424
- Muthén, B. O., & Muthén, L. K. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.
- Open Science Collaboration. (2014). The Reproducibility Project: A model of large-scale collaboration for empirical research on reproducibility. V. Stodden, F. Leisch, & R. Peng. (Eds.), *Implementing reproducible computational research (A Volume in The R Series)* (pp. 299-323). New York, NY: Taylor & Francis. doi:10.2139/ssrn.2195999
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges' (Eds.), *The handbook of research synthesis* (pp. 231-244). New York, NY: SAGE.
- Trivers, R. L. (1972). Parental investment and sexual selection. In B. G. Campbell (Ed.), *Sexual selection and the descent of man, 1871-1971* (pp. 136-179). Chicago, IL: Aldine.

- Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review, 110*, 403-421. doi:10.1037/0033-295X.110.3.403
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*, 440-463. doi:10.1037/a0018963
- Zentner, M., & Mitura, K. (2012). Stepping out of the caveman's shadow: Nations' gender gap predicts degree of sex differentiation in mate preferences. *Psychological Science, 23*(10), 1176-1185.

Author Biographies

Dylan Selterman (Ph.D., 2011, Stony Brook University) is a lecturer at the University of Maryland psychology department. His research focus is on dreams, attraction and attachment processes, and ethical behavior in close relationships. He is a contributing

editor at ScienceOfRelationships.com and the editor in chief at In-Mind Magazine.

Elizabeth Chagnon (B.Sc., 2014, University of Maryland) is a former student/research assistant with Dr. Selterman. Her research focus is on the processes of attraction and relationship development, with a special interest in LGBT populations. Currently she works at the American Institutes for Research in the Health and Social Development department, supporting the National Center for Safe Supportive Learning Environments.

Sean P. Mackinnon (Ph.D., 2012, Dalhousie University) is an instructor in the Department of Psychology and Neuroscience at Dalhousie University. His research typically focuses on the impact personality and social relationship functioning has on the development of mood problems and addictive behaviours.